

DOCUMENT RESUME

ED 346 155

TM 018 498

AUTHOR Chou, Tungshan; Huberty, Carl J.
TITLE The Robustness of the Johnson-Neyman Technique.
PUB DATE Apr 92
NOTE 30p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 20-24, 1992).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Analysis of Variance; Computer Simulation; Equations (Mathematics); *Error Patterns; Foreign Countries; *Mathematical Models; Monte Carlo Methods; *Robustness (Statistics); *Sample Size; *Statistical Distributions
IDENTIFIERS *Johnson Neyman Technique; *Type I Errors; Type III Errors

ABSTRACT

The empirical performance of the technique proposed by P. O. Johnson and J. Neyman (1936) (the JN technique) and the modification of R. F. Potthoff (1964) was studied in simulated data settings. The robustness of the two JN techniques was investigated with respect to their ability to control Type I and Type III errors. Factors manipulated in the Monte Carlo simulations were conditional distribution shape, conditional variance rate, and sample size ratio, yielding 120 simulated data conditions with 10,000 experiments for each. Results indicate that the original JN technique produces unacceptably high Type I error rates, while the modified JN technique usually controls Type I error at the nominal alpha level. The original JN technique controls Type III error at the nominal alpha level, and the modified JN technique yields a Type III error rate well below the nominal alpha level. Analysis of variance indicated that the effect of conditional distribution shape was practically negligible for both JN techniques. The effect of the conditional variance ratio was confounded by the sample size equality factor. When sample sizes were equal, the effect of the conditional variance ratio was practically negligible. However, when sample sizes were unequal, the effect of the conditional variance ratio appeared strong. There are 14 tables of analysis data, 1 figure, and a 13-item list of references. (SL)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

The Robustness of the Johnson-Neyman Technique

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☒ This document has been reproduced as
received from the person or organization
originating it.

☐ Minor changes have been made to improve
reproduction quality.

• Points of view or opinions stated in this docu-
ment do not necessarily represent official
OERI position or policy.

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

TUNGSHAN CHOU

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

Tungshan Chou
National Hualien Teachers College

Carl J Huberty
University of Georgia

Paper presented at the Annual Meeting of the
American Educational Research Association, San Francisco, April, 1992

BEST COPY AVAILABLE

ED346155

TMO18498

ABSTRACT

The purpose of this paper was to explore the empirical performance of the original Johnson-Neyman (J-N) technique and Potthoff modified J-N technique in a number of simulated data settings. The robustness properties of the two J-N techniques were investigated with respect to their ability in controlling the simultaneous error rate. The simultaneous error was defined in terms of two types of error: (1) the detection of a region of significance when in fact there is none, referred to as "Type I" error; and (2) the detected region of significance containing the covariate value at which two populations have equal expected criterion score.

Three factors were manipulated in the Monte-Carlo simulations: conditional distribution shape, conditional variance ratio, and sample size ratio. Their crossing yielded a total of 120 simulated data conditions, with 10,000 simulated experiments for each. The results indicated that the original J-N technique produced unacceptably high Type I error rates under most data conditions, whereas the modified J-N technique generally controlled the Type I error rate at the nominal alpha level. With respect to Type III error, the original J-N technique was found capable of controlling it at the nominal alpha level. For the modified J-N technique, the Type III error rate was found to be well below the nominal alpha level.

Analysis of variance (ANOVA) was used to analyze the effects of the manipulated factors on the empirical Type I and III error rates. It was found that the effect of the conditional distribution shape was practically negligible for both J-N techniques. The effect of the conditional variance ratio was confounded by the sample size equality factor. When sample sizes were equal, the effect of the conditional variance ratio was practically negligible. However, when sample sizes were unequal, the effect of the conditional variance ratio appeared strong.

INTRODUCTION

The assessment of the treatment effect for two groups in the presence of a concomitant variable has traditionally been analyzed with conventional ANCOVA under the assumption that the relationship between the concomitant variable and the criterion variable is the same for both groups. That is to say, a unit difference in the concomitant variable results in the same amount of increase or decrease in the criterion variable for both groups. The commonly referred linear model of ANCOVA is seen as,

$$Y_{ij} = \mu + \tau_j + \beta(X_{ij} - \bar{X}) + \epsilon_{ij} \quad [1]$$

where Y_{ij} is the outcome score of the i th individual in the j th treatment group ($i = 1$ to n_j), μ is the "grand mean" parameter, τ_j is the treatment effect associated with j th group, and β is the regression slope which is assumed to be the same across two populations.

Johnson and Neyman (1936) proposed a general linear hypothesis testing procedure in which the conventional ANCOVA is only a special case. Their procedure, commonly known as the Johnson-Neyman (J-N) technique, allows the within-group regression slope to vary across groups. The original J-N technique used the method of comparing two regression lines in the context of two covariates and two treatment groups. This paper looks into the robustness properties of such a procedure in the context of one covariate and two treatment groups.

Potthoff (1964) and Rogosa (1980) raised an important issue associated with the use of the J-N technique: non-simultaneous inference versus simultaneous inference. Non-simultaneous inference implies that the inferred non-zero difference between two populations is valid only for a single covariate value in the region of significance, whereas in contrast simultaneous inference implies that the inferred non-zero difference between two populations exists for all covariate values simultaneously in the region of significance. The original development of the J-N technique was said to be valid only for non-simultaneous inferences (Potthoff, 1964; Rogosa, 1980). A literature search via ERIC for Johnson-Neyman technique showed that most researchers using the original J-N technique were interested in making simultaneous inferences. Nevertheless, the exact probability of committing a simultaneous inference error by using the original J-N technique when the covariate is random cannot be theoretically determined, and consequently the extent of such inappropriateness has been unknown. Because most covariates used in educational and psychological studies are random, it appears that non-simultaneous inferences are seldom meaningful.

By using a Scheffé-like adjustment, Potthoff (1964) modified the original J-N technique so that simultaneous inference associated with such a procedure was said to be valid. The results from the literature search indicated that the use of this procedure is still rather rare. In this paper, the robustness properties of the original and modified J-N techniques were investigated with respect to their ability in controlling the simultaneous inference errors. These two procedures are briefly discussed as follows.

Original J-N technique

Johnson and Neyman (1936) fitted two separate regression lines for the two treatment groups. Let

$$\begin{aligned}\hat{Y}_{i1} &= a_0 + a_1 X_{i1}, \quad i=1 \text{ to } n_1, \text{ and} \\ \hat{Y}_{i2} &= b_0 + b_1 X_{i2}, \quad i=1 \text{ to } n_2\end{aligned}\quad [2]$$

be two separate unrestricted (i.e., allowing for unequal slopes) regression lines for Groups 1 and 2, respectively. The mean difference between two groups can be conceived as the vertical distance between these two regression lines. Let $\bar{Y}_1$ and $\bar{X}_1$ be the means of the criterion variable and covariate separately for Group 1, $\bar{Y}_2$ and $\bar{X}_2$ likewise for Group 2. The mean difference between two populations can be defined only if a covariate point is specified, and consequently it will change as a function of X . Denoted as $D(X)$,

$$D(X) = (a_0 - b_0) + (a_1 - b_1)X \quad [3]$$

Define a dummy variable T such that $T_i=1$ when the observation is from Group 1, and $T_i=0$ otherwise. Combining the two regression lines, expressions in [2] can be reparameterized as the following,

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 T_i + \hat{\beta}_2 X_i + \hat{\beta}_3 T_i X_i \quad [4]$$

where $\hat{\beta}_1 = (a_0 - b_0)$ and $\hat{\beta}_3 = (a_1 - b_1)$. The mean difference between two groups at X can thus be expressed as

$$D(X) = (a_0 - b_0) + (a_1 - b_1)X = \hat{\beta}_1 + \hat{\beta}_3 X \quad [5]$$

The variance of $D(X)$ for $X=X_p$ can be derived as

$$\begin{aligned}\text{Var}(D(X_p)) &= \text{Var}(\hat{\beta}_1 + \hat{\beta}_3 X_p) \\ &= \text{Var}(\hat{\beta}_1) + X_p^2 \text{Var}(\hat{\beta}_3) + 2X_p \cdot \text{COV}(\hat{\beta}_1, \hat{\beta}_3).\end{aligned}\quad [6]$$

A computational formula for an unbiased estimate of $\text{Var}(D(X_p))$ is given by Rogosa (1980, p. 309) as

$$S^2_{D(X_p)} = S^2 \left(\frac{1}{n_1} + \frac{1}{n_2} + \frac{(X_p - \bar{X}_1)^2}{SSX_1} + \frac{(X_p - \bar{X}_2)^2}{SSX_2} \right) \quad [7]$$

where S^2 is the error mean square calculated by summing residual sums of squares for the two within-group regressions and dividing it by the pooled error degrees of freedom:

$$S^2 = \frac{1}{N-4} \left(\sum_{i=1}^{n_1} (Y_i - a_0 - a_1 X_{i1})^2 + \sum_{i=1}^{n_2} (Y_i - b_0 - b_1 X_{i2})^2 \right) \quad [8]$$

where $N = n_1 + n_2$. The statistical significance test for $D(X_p)$ can thus be obtained by comparing the test statistic $D(X_p)/S_{D(X_p)}$ to a Student t distribution with $(N-4)$ degrees of freedom. To construct a $100(1-\alpha)\%$ confidence interval for $D(X)$, we have

$$D(X) \pm t_{\alpha/2, N-4} \sqrt{\text{Var}(\hat{\beta}_1) + X^2 \text{Var}(\hat{\beta}_3) + 2X \cdot \text{COV}(\hat{\beta}_1, \hat{\beta}_3)}. \quad [9]$$

The original J-N technique finds the set of X values such that the above confidence interval does not contain zero. This set of values is called the region of significance, denoted by R . Thus, R is composed of all X values that satisfy the inequality

$$[D(X)]^2 > F_{\alpha,1,N-4} \left(\hat{\text{Var}}(\hat{\beta}_1) + X^2 \hat{\text{Var}}(\hat{\beta}_3) + 2X \cdot \hat{\text{Cov}}(\hat{\beta}_1, \hat{\beta}_3) \right) \quad [10]$$

Expanding the above inequality, one finds that the bounds for R are the roots of a quadratic equation in X :

$$\begin{aligned} & [\hat{\beta}_3^2 - F_{\alpha,1,N-4} \hat{\text{Var}}(\hat{\beta}_3)] X^2 + \\ & 2[\hat{\beta}_1 \hat{\beta}_3 - F_{\alpha,1,N-4} \hat{\text{Cov}}(\hat{\beta}_1, \hat{\beta}_3)] X + \\ & [\hat{\beta}_1^2 - F_{\alpha,1,N-4} \hat{\text{Var}}(\hat{\beta}_1)] = 0. \end{aligned} \quad [11]$$

Equation [11] takes on the form $AX^2 + 2BX + C = 0$. If R exists, the two roots of this equation are

$$\frac{-B \pm \sqrt{B^2 - AC}}{A} \quad [12]$$

Let X_+ denote the larger root and X_- the smaller root, Rogosa (1981) showed that R is composed of the X values: $X > X_+$ or $X < X_-$, when $\hat{\beta}_3^2 / \hat{\text{Var}}(\hat{\beta}_3) > F_{\alpha,1,N-4}$; and R is composed of the X values: $X_- < X < X_+$, when $\hat{\beta}_3^2 / \hat{\text{Var}}(\hat{\beta}_3) < F_{\alpha,1,N-4}$. Notice that in order to have real solutions, the quantity under the radical of [12], $(B^2 - AC)$, has to be non-negative.

Because the confidence interval in expression [9] is valid only for a specific X value, the derived region of significance is non-simultaneous. Potthoff (1964) and Rogosa (1980) have argued that the probability statement for such a region has been incorrectly interpreted as if treatment mean differences are present for all values in R . Statements about the vertical distances between two regression lines based on the non-simultaneous region are valid only for a single specific value of X , not for all values in R . Since the exact type I error probability of making a simultaneous inference based on the non-simultaneous region of significance is not known, the extent of the seriousness of making such wrong inferences cannot be determined.

The J-N technique has been presented in several statistical methods books as an alternative to ANCOVA when the test of slopes homogeneity is rejected (Huitema, 1980; Pedhazur, 1982). The test of slope homogeneity is rejected when $\hat{\beta}_3^2 / \hat{\text{Var}}(\hat{\beta}_3) > F_{\alpha,1,N-4}$. Therefore, the J-N technique is presented in these two books as to yield a region of nonsignificance between X_- and X_+ . Chou and Wang (1992) investigated the necessity of such a test preceding the J-N technique, interested readers may want to refer to their paper.

Modified Johnson-Neyman Technique

Potthoff (1964) modified the original J-N technique so that one may make statements about the mean differences between two populations over the entire region of significance. Denote the simultaneous region of significance by R' ; we can state that two population means are different simultaneously for all points contained in R' . Potthoff (1964) defined such a region of significance (for $\alpha = .05$) as the following:

In the long run, not more than 5 per cent of such regions which are calculated will contain any points at all for which the two groups are equal in expected criterion score. (p. 244)

The simultaneous region of significance can be derived by constructing a $100(1-\alpha)\%$ simultaneous confidence band for the differences between two groups. The difference between two groups can be expressed as a function of X , denoted by a straight line $Z = D(X) = \hat{\beta}_1 + \hat{\beta}_3 X$. Potthoff (1964) applied the Working-Hotelling band (Miller, 1966; p. 111) to modify the original J-N method in deriving a simultaneous region of significance. The region is enclosed by the upper and lower confidence limits,

$$D(X) \pm \sqrt{2F_{\alpha,2,N-4} \left(\hat{\text{Var}}(\hat{\beta}_1) + X^2 \hat{\text{Var}}(\hat{\beta}_3) + 2X \hat{\text{Cov}}(\hat{\beta}_1, \hat{\beta}_3) \right)} \quad [13]$$

Expressions [13] and [9] are different only in the numerator degrees of freedom of the F distribution and a constant multiplier: [9] uses $t_{\alpha/2, N-4}^2 = F_{\alpha,1,N-4}$ whereas [13] uses $2F_{\alpha,2,N-4}$. The bounds of R' in [13] are thus the roots of a quadratic equation,

$$\begin{aligned} & [\hat{\beta}_3^2 - 2F_{\alpha,2,N-4} \hat{\text{Var}}(\hat{\beta}_3)] X^2 + 2[\hat{\beta}_1 \hat{\beta}_3 - 2F_{\alpha,2,N-4} \hat{\text{Cov}}(\hat{\beta}_1, \hat{\beta}_3)] X + \\ & [\hat{\beta}_1^2 - 2F_{\alpha,2,N-4} \hat{\text{Var}}(\hat{\beta}_1)] = 0. \end{aligned} \quad [14]$$

Substituting proper A , B , C values into [12], two roots are obtained. Denote X_+ for the larger root and X_- for the smaller root. When $\hat{\beta}_3^2 / \hat{\text{Var}}(\hat{\beta}_3) > 2F_{\alpha,2,N-4}$, R' is composed of the X values: $X > X_+$ or $X < X_-$. When $\hat{\beta}_3^2 / \hat{\text{Var}}(\hat{\beta}_3) < 2F_{\alpha,2,N-4}$, R' is composed of the X values: $X_- < X < X_+$. Because $2F_{\alpha,2,N-4} > F_{\alpha,1,N-4}$, R' is always a subset of R .

Simultaneous Inference Error

Two types of simultaneous inference error are conceivable with both J-N techniques: (1) the detection of a region of significance when in fact there is none, and (2) the detected region of significance contains the point where the two population means are equal. The former type of error may be committed under the situation where the two populations have identical regression lines. It may be assessed by the percent of the computed regions of significance under the simulated setting of two identical regression lines. The latter type of error may be committed under the situation when two population regression lines are not identical. In fact, any two non-identical linear population regression lines intersect at only one covariate point, at which the two population means are equal. Consequently, this type of error may be assessed by the percent of times the detected region of significance contains this intersection point. In this paper, the former type of error will be referred to as the "Type I error," whereas the latter type will be referred to as the "Type III error."

MONTE CARLO SIMULATION PROCEDURES

The empirical performances of the two J-N techniques were assessed under two settings of the regression parameters: complete null (i.e., equal slopes and intercepts) and heterogeneity of slopes (i.e., intersecting

regression lines). "Type I" error and "Type III" error were investigated in their respective setting. Three factors were manipulated under each setting: (a) conditional distribution shape, (b) conditional variance ratio, and (c) sample size ratio. The first two factors will be described later in this section. Of the five sample size ratios, three were equal: (10:10), (20:20), and (30:30); and two were unequal: (10:20), and (10:30). The three factors were completely crossed. There were $8 \times 3 \times 5 = 120$ simulation conditions with 10,000 simulated experiments in each simulation condition. With two settings of the regression parameters, the total number of simulated experiments in the study was 2,400,000. All tests of significance and the computation of the regions of significance were done at a nominal alpha level of .05.

Data Simulation

The data for the two groups were simulated according to the following regression model equations:

$$\begin{aligned} Y_{i1} &= \alpha_0 + \alpha_1 X_{i1} + \epsilon_{i1}, \text{ for group 1, and} \\ Y_{i2} &= \beta_0 + \beta_1 X_{i2} + \epsilon_{i2}, \text{ for group 2.} \end{aligned} \quad [15]$$

The X variable was a random-valued covariate generated by converting a standard normal variable Z_1 to a normal random variable with a mean of 20 and variance of 9 for both groups. The values of 20 and 9 were selected such that the observed covariate score range would fall mostly between 10 and 30, which is typical of many covariates in psychology and education. The variable Z_1 was generated by the RANNOR function of the Statistical Analysis System (SAS, 1985). The ϵ_i 's were the random error components from a desired distribution with a mean of 0 and selected variance values. The selections of the population regression parameters were presented in Table 1. The selections of the distribution shape and variance for the ϵ_i 's will be described in later sections.

Shapes of the Conditional Distributions

The eight shapes of the conditional distributions adopted here were taken from Olejnik and Algina's (1987) study of the effect of heteroscedastic conditional variance on the Type I error and statistical power of the conventional ANCOVA. They are shown in Table 2. These eight distributional shapes were chosen because it is believed that they are representative of many educational and psychological data distributions. These eight distributions were simulated by choosing the appropriate a , b , c , and d values in a table provided by Fleishman (1978) for

$$E = a + bZ_2 + cZ_2^2 + dZ_2^3,$$

where Z_2 is a standard normal variable. The resulting E is distributed with a mean of 0 and variance of 1. Then, E was multiplied by an appropriate constant (the positive square root of the desired conditional variance) to change its variance to a preset value for each of the two groups. The generated E was a random variable having the desired distribution shape, a mean of 0, and desired variance value. The E 's were used as the ϵ_i terms in expressions [15] for generating the outcome scores.

Conditional Variance Ratio

Three conditional variance ratios were considered in the current study: (36:36), (18:54), and (54:18). The values were chosen such that the average of the paired conditional variances was 36 (for the convenience of setting the effect size to be 2/3 in parallel lines setting) and that the larger conditional variance could be associated with either larger or smaller sample size in the pair. These ratios were used to represent conditional variance homogeneity and heterogeneity conditions that are likely to be encountered in educational and psychological data.

Statistical Analysis of Simulation Results

Three factors were investigated for their effects on the empirical error rates: the shape of the conditional distributions, conditional variance ratio, and sample size equality. The sample size equality factor was coded as 1 if two groups had the same sample sizes (i.e., 10:10, 20:20, and 30:30), or 0 if two groups had unequal sample sizes (i.e., 10:20, and 10:30).

Harvell (1991) suggested a need for using research designs to detect key patterns in Monte Carlo studies. We employed a three-way fixed-effect analysis of variance (ANOVA) to analyze the effects of these three factors on the empirical type I & III error rates. The schematic of for the 8*3*2 factorial design is shown in Figure 1.

To obtain an estimate of the empirical error rate and its variability, a 95% confidence interval for the mean of the respective error rate was constructed for each level of the factors in concern. The observed mean error rate was judged to be within statistical tolerance of the nominal alpha level if the 95% confidence interval contained the nominal value of .05.

SIMULATION RESULTS

Analyses of Type I error Rate

The ANOVA results for the original J-N technique are reported in Table 3. The eta-squared values (η^2) as reported here are the effect size indicators for the associated effects in relation to the error sum of squares. It is bounded by 0 and 1 with higher values indicating stronger relationship with the empirical Type I error rate.

There was no significant three-way interaction effect ($F(14,72) = .16, p = .9998, \eta^2 = .030$). The only significant two-way interaction was found between sample size equality (EQ) and conditional variance ratio (VAR) ($F(2,72) = 503.93, p < .0001, \eta^2 = .933$), which cautions against the direct interpretation of either main effect. The mean Type I error rate appeared not affected by the shape of the conditional distributions (DS) ($F(7,72) = .24, p = .9733, \eta^2 = .023$). The overall R-squared value of .960 indicated that most of the variation (approximately 96%) in the observed Type I error rates could be attributed to the effects as specified in the ANOVA model.

Based on the ANOVA results, the effect of DS on the Type I error rate was minimal. Also because all the interaction effects associated with DS

were minimal, it was deemed appropriate that the summary statistics for the eight levels of DS, collapsed over EQ and VAR, be reported as in Table 4. Here it shows that the means of the empirical Type I error rates for all eight conditional distribution shapes were well beyond the nominal value of .05. The lower limit of the 95% confidence interval (LL) was approximately .12 for all eight conditional distributions. The differences in the means of the Type I error rates for the eight conditional distributions were very slight (in the thousandths).

The effect of the VAR*EQ interaction was seen as strong. It was therefore deemed appropriate that the summary statistics for the three levels of VAR by two levels of EQ, collapsed over DS, be shown as in Table 5. When sample sizes were equal (EQ = 'E'), it appeared that the conditional variance ratio did not affect the mean Type I error rate. When equal sample sizes and equal conditional variance ratios were considered (EQ = "E" and VAR = "2"), the mean Type I error rate was .1417. When sample sizes were equal, the mean Type I error rate for VAR = "1" and "3" was very similar to that for VAR = "2" (the differences were in thousandths). However, when sample sizes were not equal (EQ = 'U'), the mean Type I error rate changed greatly depending on the conditional variance ratio. When the smaller sample size group was associated with the smaller conditional variance (EQ = "U" and VAR = "1"), the observed mean Type I error rate was .0593. When unequal sample sizes were associated with equal conditional variance ratio (EQ = 'U' and VAR = "2"), the mean empirical Type I error rate was about the same level as those obtained when EQ = "E". When the smaller sample size group was associated with the larger conditional variance (EQ = "U" and VAR = "3"), the mean empirical Type I error rate was as high as .2844.

The ANOVA results for the modified J-N technique are reported in Table 6. These results are quite comparable to those for the original J-N technique. The only significant two-way interaction was found between EQ and VAR ($F(2,72) = 298.44$, $p < .0001$, $\eta^2 = .892$), which cautions against the direct interpretation of either main effect. The mean Type I error rate appeared not affected by DS ($F(7,72) = .28$, $p = .9613$, $\eta^2 = .026$). The overall R-squared value of .937 indicated that most of the variation (approximately 94%) in the observed Type I error rates could be attributed to the effects as specified in the ANOVA model.

Based on the ANOVA results, it was apparent that the effect of DS on the Type I error rate was minimal (the p-value was almost approximatedly 1.0). Also because all the interaction terms associated with DS were minimal, it was therefore deemed appropriate for the sake of simplicity that the summary statistics for the eight levels of DS, collapsed over EQ and VAR, be reported as in Table 7. Here it shows that the means of the empirical Type I error rates for all eight conditional distributions were approximately .06. The largest mean Type I error rate was .0630 when the conditional distribution shape was highly skewed and moderately leptokurtic (DS = "G"). The 95% confidence intervals (as shown by LL and UL) for the mean Type I error rates for all eight conditional distributions contained the nominal value of .05. The interquartile range ($Q3 - Q1$) of the empirical Type I error rate for all eight conditional distribution shapes was less than .02. The largest interquartile range was .0139 ($= .0621 - .0482$) when DS = "G". The low variability of the empirical Type I error rate as indicated by the narrow interquartile range showed that the modified J-N technique was able to maintain a high precision of controlling

the Type I error rate at the nominal level for all eight conditional distribution shapes.

The effect of the VAR*EQ interaction was strong. It was deemed appropriate that the summary statistics for the three levels of VAR by two levels of EQ, collapsed over DS, be reported as in Table 8. When sample sizes were equal (EQ = 'E'), the mean empirical Type I error rates for VAR = "1" (18:54), "2" (36:36), and "3" (54:18) were .0547, .0492, and .0538, respectively. The differences between UL and LL were within .01. However, when sample sizes were not equal (EQ = 'U'), the mean Type I error rate changed greatly depending on the conditional variance ratio. When the smaller sample size group was associated with the smaller conditional variance (EQ = "U" and VAR = "1"), the observed mean Type I error rate was .0161. When unequal sample sizes were associated with unequal conditional variance ratio (EQ = 'U' and VAR = "2"), the mean empirical Type I error rate was about the same level as those obtained when EQ = "E". When the smaller sample size group was associated with the larger conditional variance (EQ = "U" and VAR = "3"), the mean empirical Type I error rate was as high as .1452.

Analyses of Type III error Rate

The ANOVA results for the original J-N technique are reported in Table 9. There was no significant three-way interaction effect ($F(14,72) = .08$, $p = 1.0000$, $\eta^2 = .005$). The only significant two-way interaction was found between sample size equality (EQ) and conditional variance ratio (VAR) ($F(2,72) = 327.35$, $p = .0000$, $\eta^2 = .740$), which cautions against the direct interpretation of either main effect. The mean Type III error rate appeared not affected by the conditional distribution shape (DS) ($F(7,72) = .58$, $p = .7737$, $\eta^2 = .017$). The overall R-squared value of .942 indicated that most of the variation (approximately 94%) in the observed Type III error rates could be attributed to the effects as specified in the ANOVA model.

Based on the ANOVA results, the effect of DS on the Type III error rate was minimal. Also because all interaction effects associated with DS were minimal, it was therefore deemed appropriate that the summary statistics for the eight levels of DS, collapsed over EQ and VAR, are reported in Table 10. The means of the empirical Type III error rates for all eight conditional distributions were approximately .057. The largest empirical mean Type III error rate was .0609 when DS = "G" (highly skewed and moderately leptokurtic). The 95% confidence intervals for the mean Type III error rates for all eight conditional distributions contained the nominal value of .05. The interquartile range ($Q3 - Q1$) of the empirical Type III error rate for all eight conditional distribution shapes was less than .02. The largest interquartile range was .0171 ($= .0657 - .0486$) when DS = "G". The low variability of the empirical Type III error rate as indicated by the narrow interquartile range showed that the original J-N technique was able to maintain a high precision of controlling the Type III error rate at the nominal level for all eight conditional distribution shapes.

The effect of the VAR*EQ interaction was seen as strong. The summary statistics for the three levels of VAR by two levels of EQ, collapsed over DS, are shown in Table 11. When sample sizes were equal (EQ = 'E'), the mean empirical Type III error rates for VAR = "1" (18:54), "2" (36:36), and "3" (54:18) were .0537, .0503, and .0543, respectively. The differences

between UL and LL were approximately within .01. However, when sample sizes were not equal (EQ = 'U'), the mean Type III error rate changed greatly depending on the conditional variance ratio. When the smaller sample size group was associated with the smaller conditional variance (EQ = "U" and VAR = "1"), the observed mean Type III error rate was .0186. When unequal sample sizes were associated with equal conditional variance ratio (EQ = 'U' and VAR = "2"), the mean empirical Type III error rate was .503, which is about the same level as those obtained when EQ = "E". When the smaller sample size group was associated with the larger conditional variance (EQ = "U" and VAR = "3"), the mean empirical Type III error rate was as high as .1203.

The ANOVA results for the modified J-N technique are reported in Table 12. These results are quite comparable to the ANOVA results on the Type III error rates for the original J-N technique. There was no significant three-way interaction effect ($F(14,72) = .16$, $p = .9998$, $\eta^2 = .030$). The only significant two-way interaction was found between EQ and VAR ($F(2,72) = 201.17$, $p < .0000$, $\eta^2 = .848$), which cautions against the direct interpretation of either main effect. The mean Type III error rate appeared not affected by DS ($F(7,72) = .78$, $p = .6093$, $\eta^2 = .070$). The overall R-squared value of .913162 indicated that most of the variation (approximately 91%) in the observed Type III error rates could be attributed to the effects as specified in the ANOVA model.

As seen from the ANOVA results, the effect of DS on the Type III error rate was minimal. The summary statistics for the eight levels of DS, collapsed over EQ and VAR, are reported in Table 13. The means of the empirical Type III error rates for all eight conditional distributions were approximately .020. The largest mean empirical Type III error rate was .0222 for DS = "G" (highly skewed and moderately leptokurtic). The upper limit of the 95% confidence intervals for the mean Type III error rate for all eight conditional distributions was below .03. The interquartile range of the empirical Type III error rate for all eight conditional distribution shapes was less than .02. The largest interquartile ($Q3 - Q1$) range was .0147 ($= .0274 - .0127$) for DS = "G". The low variability of the empirical Type III error rate as indicated by the narrow interquartile range showed that the modified J-N technique was able to maintain a high precision of controlling the Type III error rate well below the nominal value of .05.

The effect of the VAR*EQ interaction was seen as strong. The summary statistics for the three levels of VAR by two levels of EQ, collapsed over DS, are shown in Table 14. When sample sizes were equal (EQ = 'E'), the mean empirical Type III error rates for VAR = "1" (18:54), "2" (36:36), and "3" (54:18) were .0178, .0153, and .0176, respectively. The differences between UL and LL were within .003. However, when sample sizes were not equal (EQ = 'U'), the mean Type III error rate changed greatly depending on the conditional variance ratio. When the smaller sample size group was associated with the smaller conditional variance (EQ = "U" and VAR = "1"), the observed mean Type III error rate was .0046. When unequal sample sizes were associated with equal conditional variance ratio (EQ = 'U' and VAR = "2"), the mean empirical Type III error rate was about the same level as those obtained when EQ = "E". When the smaller sample size group was associated with the larger conditional variance (EQ = "U" and VAR = "3"), the mean empirical Type III error rate was as high as .0536.

CONCLUSIONS

Both J-N techniques have similar robustness (as measured by the Type I error rate and Type III error rate) properties concerning the effects of the conditional distribution shape, conditional variance ratio, and sample size equality factor. The original J-N technique however, yielded unacceptably high Type I error rate under virtually all data conditions whereas the modified J-N technique generally controlled the Type I error rate at the nominal level within statistical tolerance. These findings were similar to those reported by Shields (1978). Both techniques also showed low variability for both the mean type I and Type III error rates, which indicated high consistency of the empirical error rates.

For both techniques, the effect of the conditional distribution shape on both error rates was minimal. The main effect of the conditional variance ratio was confounded by the sample size equality factor. When sample sizes were unequal, both techniques tend to be conservative if the smaller conditional variance was associated with the smaller sample size group. When unequal sample sizes were associated with equal conditional variance ratios, the modified J-N technique still controlled the Type I error rate at the nominal α level. The modified J-N technique tends to be liberal (produced unacceptably high Type I error rate) if the larger conditional variance was associated with the smaller sample size group. The Type III error rates produced by the original J-N technique under most data conditions were at the nominal level within statistical tolerance. The Type III error rate for the modified J-N technique was generally well below the nominal α level of .05 (approximately .02).

DISCUSSION

This study showed that robustness properties of both J-N techniques were similar to those of the conventional ANCOVA. With respect to making simultaneous inferences when the slopes are different, the Type I error was redefined for the original and modified J-N techniques as the detection of a region of significance under the complete null situation. In the current study, the original J-N technique was found to yield an unacceptably high Type I error rate in the complete null data condition even when all parametric assumptions are met. This may be expected because the development of the significance region for the original J-N technique was based upon the statistical significance test at a single covariate point. The results of the current study showed that mechanical applications of the original J-N technique as suggested by Rogosa (1980, p. 313) may likely obtain a region of significance when in fact there is none.

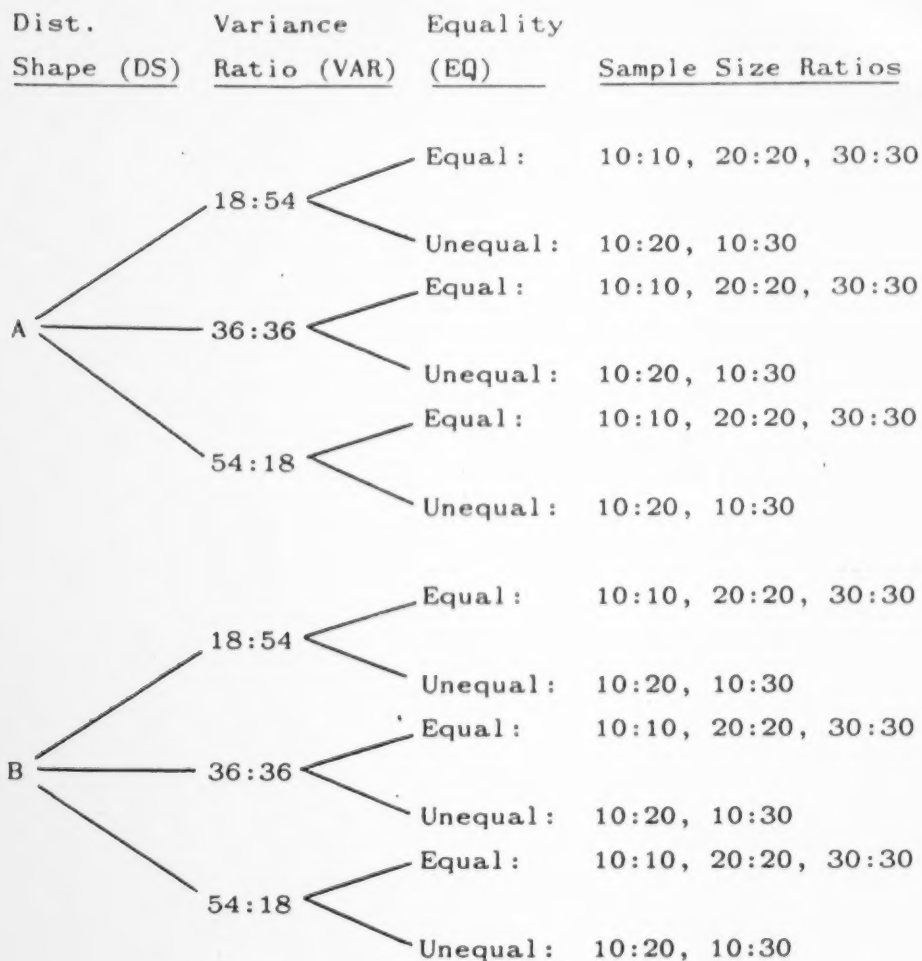
Fortunately, the Type III error rate yielded by the original J-N technique was found in the current study to approximate the nominal α level. This finding is encouraging because it indicates that the original J-N technique may be appropriate for making simultaneous inferences when slopes are in fact unequal (No Type I error would be committed if two lines are non-parallel). Nevertheless, researchers are often unaware of the population slope heterogeneity and rely on statistical testing to conclude that there is indeed a significant difference in slopes. Results of the current study seem to support the notion that the common practice as suggested by Huitema (1980, p. 270) for using the original J-N technique only after the detection of a significant covariate-treatment interaction

is appropriate. Such a practice appears theoretically feasible because a region of significance will always exist if the population slopes are not equal. Chou and Wang (1992) investigated the effect of the extent of slope heterogeneity on the Type III error rate of the original J-N technique under the condition when the slope homogeneity assumption is rejected and under the condition when the slope homogeneity assumption is not rejected. Interested readers may refer to their paper.

The conservativeness of the Type III error rate yielded by the modified J-N technique is not surprising due to the Scheffé-like adjustment in the computation of the significance region. Potthoff (1964) recognized such conservativeness and suggested the adoption of a more liberal α level for the computation of the region of significance. The present study showed that the original J-N technique was actually capable of controlling the Type III error rate at the nominal α level under the slope heterogeneity situation. If the Type I error no longer presents a threat to the researcher as in the case of rejecting the slope homogeneity assumption, it should be the choice of analysis for the sake of obtaining a wider significance region.

REFERENCES

- Chon, T. , & Wang, L. S. (1992, April). Making Simultaneous Inferences Using Johnson-Neyman Technique. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco.
- Fleishman, A. I. (1978). A method for simulating non-normal distributions. *Psychometrika*, 43, 521-531.
- Harvell, M. R. (1991, April). Communicating Monte Carlo Results in Educational and Psychological Research. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago.
- Huitema, B. E. (1980). The Analysis of Covariance and Alternatives. New York: Wiley.
- Johnson, P. O. , & Neyman, J. (1936). Tests of certain linear hypotheses and their applications to some educational problems. Statistical Research Memoirs, 1, 57-93.
- Miller, R. G. (1966). Simultaneous Statistical Inference, New York: McGraw-Hill.
- Potthoff, R. F. (1964). On the Johnson-Neyman technique and some extensions thereof. *Psychometrika*, 29, 241-256.
- Rogosa, D. (1978). Some results for the Johnson-Neyman technique. Dissertation Abstracts International, 38, 5366A. (University Microfilms No. 78-02,225).
- Rogosa, D. (1980). Comparing nonparallel regression lines. *Psychological Bulletin*, 88, 307-321.
- Rogosa, D. (1981). On the relationship between the Johnson-Neyman region of significance and statistical tests of parallel within-group regressions. *Educational and Psychological Measurement*, 41, 73-84.
- SAS Institute Inc. (1982). SAS user's guide: Statistics, (1982 edition), Cary, NC.
- SAS Institute Inc. (1985). SAS user's guide: Basics, (5th edition), Cary, NC.
- Shields, J. L. (1978). An empirical investigation of the effect of heteroscedasticity and heterogeneity of variance on the analysis of covariance and the Johnson-Neyman technique. Technical paper 292, U.S. Army Research Institute for the Behavioral and Social Sciences, Alexandria, Virginia.



. Same structure as above from here on.

.

.

.

G

N

Figure 1

A Schematic for the Three-Way ANOVA on Type-I and Type III
Error Rates

Table 1

Regression Coefficients and the Strength of the Criterion-
Covariate Relationship in the TWO Simulation Settings

Values of (α_0, α_1) for group 1 and (β_0, β_1) for group 2

	Null	Intersecting
Group 1:	10, 2	20, 2
Group 2:	10, 2	100, -2

Values of ρ^2 for Each Value of $\text{Var}(Y|X)$

$\text{Var}(Y X)$	18	36	54
ρ^2	.667	.500	.400

Table 2

Shapes of the Conditional Distributions

Distribution Shape	Skewness	Kurtosis
Platykurtic (A)	0	-1.00
Moderately Skewed (B)	.75	0
Slightly Leptokurtic (C)	0	1.00
Moderately Leptokurtic (D)	0	3.75
Slightly Skewed, and Moderately Leptokurtic (E)	.50	3.75
Moderately Skewed, and Moderately Leptokurtic (F)	1.00	3.75
Highly Skewed, and Moderately Leptokurtic (G)	1.75	3.75
Normal (N)	0	0

Note: This table was taken from Olejnik and Algina (1987).

Table 3.

Analysis of Variance on Type I Error Rates for the Original J-N
Technique

Source	df	Sum of Squares	F	Prob > F	Eta-squared
DS	7	.000417	.24	.9733	.023
VAR	2	.249512	505.69	.0000	.934
EQ	1	.007853	31.83	.0000	.307
DS=EQ	7	.000091	.05	.9998	.005
DS=VAR	14	.000807	.23	.9979	.043
EQ=VAR	2	.248646	503.93	.0000	.933
DS=EQ= VAR	14	.000551	.16	.9998	.030
Error	72	.017763			

R-squared = .959873.

Note. DS = Shape of conditional distributions; VAR = Conditional variance ratio; EQ = Sample size equality; Eta-squared = $SSH / (SSH + SSE)$. The sums of squares were obtained using the Type III sums of squares as defined in SAS PROC GLM procedure.

Table 4.

Summary Statistics on the Type I Error Rates Yielded by the
Original J-N Technique for Eight Shapes of the Conditional
Distribution

DS	n	Mean	LL	UL	Max	Q3	Med	Q1	Min
A	15	.1532	.1198	.1866	.3238	.1519	.1473	.1417	.0406
B	15	.1527	.1202	.1852	.3200	.1496	.1451	.1400	.0493
C	15	.1511	.1192	.1831	.3122	.1487	.1442	.1406	.0400
D	15	.1493	.1197	.1790	.2996	.1477	.1428	.1382	.0467
E	15	.1503	.1195	.1811	.3077	.1484	.1440	.1387	.0493
F	15	.1517	.1208	.1826	.3090	.1487	.1460	.1416	.0494
G	15	.1560	.1242	.1878	.3224	.1570	.1488	.1377	.0521
N	15	.1526	.1194	.1859	.3192	.1500	.1447	.1378	.0441

Note. DS = Conditional distribution shape; n = Number of observations; LL = Lower limit of the 95% confidence interval for the mean; UL = Upper limit of the 95% confidence interval for the mean; Max = Maximum value; Q3 = 75th percentile; Med = Median; Q1 = 25th percentile; Min = Minimum value.

Table 5

Summary Statistics on the Type I Error Rates Yielded by the
Original J-N Technique for Three Conditional Variance Ratios
by Equal and Unequal Sample Sizes

EQ	VAR	n	Mean	LL	UL	Max	Q3	Med	Q1	Min
E	1	24	.1478	.1460	.1496	.1599	.1497	.1471	.1446	.1406
E	2	24	.1417	.1399	.1436	.1526	.1450	.1412	.1377	.1354
E	3	24	.1470	.1456	.1485	.1548	.1492	.1463	.1444	.1399
U	1	16	.0593	.0523	.0663	.0844	.0710	.0590	.0473	.0400
U	2	16	.1423	.1409	.1438	.1477	.1440	.1418	.1401	.1377
U	3	16	.2844	.2689	.3000	.3238	.3174	.2825	.2546	.2434

Note. EQ = Sample size equality (E = Equal, U = Unequal); VAR = Conditional variance ratio (1 = 18:54, 2 = 36:36, 3 = 54:18); n = Number of observations; LL = Lower limit of the 95% confidence interval for the mean; UL = Upper limit of the 95% confidence interval for the mean; Max = Maximum value; Q3 = 75th percentile; Med = Median; Q1 = 25th percentile; Min = Minimum value.

Table 6

Analysis of Variance on Type I Error Rates for the Modified J-N
Technique

Source	df	Sum of Squares	F	Prob > F	Eta-squared
DS	7	.000275	.28	.9613	.026
VAR	2	.086502	304.01	.0000	.894
EQ	1	.009622	67.62	.0000	.484
DS=EQ	7	.000069	.07	.9994	.007
DS=VAR	14	.000436	.22	.9986	.041
EQ=VAR	2	.084917	298.44	.0000	.892
DS=EQ= VAR	14	.000295	.15	.9998	.028
Error	72	.010243			

R-squared = .937492.

Note. DS = Shape of conditional distributions; VAR = Conditional variance ratio; EQ = Sample size equality; Eta-squared = $SSH / (SSH + SSE)$. The sums of squares were obtained using the Type III sums of squares as defined in SAS PROC GLM procedure.

Table 7

Summary Statistics on the Type I Error Rates Yielded by the
Modified J-N Technique for Eight Shapes of the Conditional
Distribution

DS	n	Mean	LL	UL	Max	Q3	Med	Q1	Min
A	15	.0610	.0406	.0814	.1768	.0573	.0526	.0492	.0078
B	15	.0605	.0402	.0808	.1758	.0584	.0503	.0485	.0128
C	15	.0592	.0395	.0790	.1699	.0552	.0508	.0495	.0094
D	15	.0582	.0404	.0760	.1585	.0559	.0516	.0483	.0113
E	15	.0577	.0386	.0768	.1686	.0538	.0502	.0463	.0108
F	15	.0598	.0419	.0778	.1589	.0573	.0530	.0507	.0115
G	15	.0630	.0439	.0821	.1730	.0621	.0549	.0482	.0165
N	15	.0598	.0396	.0799	.1724	.0579	.0500	.0470	.0099

Note. DS = Conditional distribution shape; n = Number of observations; LL = Lower limit of the 95% confidence interval for the mean; UL = Upper limit of the 95% confidence interval for the mean; Max = Maximum value; Q3 = 75th percentile; Med = Median; Q1 = 25th percentile; Min = Minimum value.

Table 8

Summary Statistics on the Type I Error Rates Yielded by the
Modified J-N Technique for Three Conditional Variance Ratios
by Equal and Unequal Sample Sizes

EQ	VAR	n	Mean	LL	UL	Max	Q3	Med	Q1	Min
E	1	24	.0547	.0528	.0566	.0672	.0583	.0547	.0515	.0464
E	2	24	.0492	.0483	.0501	.0530	.0508	.0491	.0477	.0452
E	3	24	.0538	.0524	.0552	.0621	.0562	.0535	.0512	.0469
U	1	16	.0161	.0132	.0191	.0302	.0200	.0175	.0109	.0078
U	2	16	.0512	.0500	.0525	.0559	.0535	.0503	.0493	.0472
U	3	16	.1452	.1327	.1577	.1768	.1718	.1424	.1227	.1137

Note. EQ = Sample size equality (E = Equal, U = Unequal); VAR = Conditional variance ratio (1 = 18:54, 2 = 36:36, 3 = 54:18); n = Number of observations; LL = Lower limit of the 95% confidence interval for the mean; UL = Upper limit of the 95% confidence interval for the mean; Max = Maximum value; Q3 = 75th percentile; Med = Median; Q1 = 25th percentile; Min = Minimum value.

Table 9

Analysis of Variance on Type III Error Rates for the Original J-N
Technique

Source	df	Sum of Squares	F	Prob > F	Eta-squared
DS	7	.000311	.58	.7737	.017
VAR	2	.053559	347.13	.0000	.751
EQ	1	.003063	39.71	.0000	.147
DS-EQ	7	.000071	.13	.9956	.004
DS-VAR	14	.000437	.41	.9689	.024
EQ-VAR	2	.050506	327.35	.0000	.740
DS-EQ- VAR	14	.000090	.08	1.0000	.005
Error	72	.017763			

R-squared = .942377.

Note. DS = Shape of conditional distributions; VAR = Conditional variance ratio; EQ = Sample size equality; Eta-squared = $SSH / (SSH + SSE)$. The sums of squares were obtained using the Type III sums of squares as defined in SAS PROC GLM procedure.

Table 10

Summary Statistics on the Type III Error Rates Yielded by the
Original J-N Technique for Eight Shapes of the Conditional
Distribution

DS	n	Mean	LL	UL	Max	Q3	Med	Q1	Min
A	15	.0568	.0418	.0717	.1381	.0607	.0510	.0484	.0137
B	15	.0579	.0426	.0732	.1398	.0631	.0517	.0491	.0126
C	15	.0561	.0415	.0707	.1329	.0533	.0520	.0499	.0125
D	15	.0555	.0414	.0697	.1319	.0544	.0504	.0495	.0143
E	15	.0548	.0408	.0688	.1300	.0527	.0505	.0485	.0140
F	15	.0561	.0409	.0712	.1414	.0551	.0501	.0470	.0156
G	15	.0609	.0457	.0760	.1414	.0657	.0545	.0486	.0200
N	15	.0569	.0417	.0721	.1386	.0549	.0522	.0486	.0103

Note. DS = Conditional distribution shape; n = Number of observations; LL = Lower limit of the 95% confidence interval for the mean; UL = Upper limit of the 95% confidence interval for the mean; Max = Maximum value; Q3 = 75th percentile; Med = Median; Q1 = 25th percentile; Min = Minimum value.

Table 11

Summary Statistics on the Type III Error Rates Yielded by the
Original J-N Technique for Three Conditional Variance Ratios
by Equal and Unequal Sample Sizes

EQ	VAR	n	Mean	LL	UL	Max	Q3	Med	Q1	Min
E	1	24	.0537	.0516	.0558	.0688	.0557	.0526	.0495	.0473
E	2	24	.0503	.0493	.0512	.0549	.0521	.0500	.0484	.0465
E	3	24	.0543	.0522	.0564	.0672	.0565	.0519	.0505	.0480
U	1	16	.0186	.0160	.0213	.0301	.0229	.0202	.0138	.0103
U	2	16	.0503	.0491	.0514	.0536	.0522	.0504	.0491	.0444
U	3	16	.1203	.1117	.1288	.1414	.1385	.1202	.1027	.0993

Note. EQ = Sample size equality (E, = Equal, U = Unequal); VAR = Conditional variance ratio (1 = 18:54, 2 = 36:36, 3 = 54:18); n = Number of observations; LL = Lower limit of the 95% confidence interval for the mean; UL = Upper limit of the 95% confidence interval for the mean; Max = Maximum value; Q3 = 75th percentile; Med = Median; Q1 = 25th percentile; Min = Minimum value.

Table 12

Analysis of Variance on Type III Error Rates for the Modified J-N
Technique

Source	df	Sum of Squares	F	Prob > F	Eta-squared
DS	7	.000169	.78	.6093	.070
VAR	2	.013075	210.61	.0000	.854
EQ	1	.001695	54.62	.0000	.431
DS=EQ	7	.000015	.07	.9994	.007
DS=VAR	14	.000177	.41	.9677	.073
EQ=VAR	2	.012489	201.17	.0000	.848
DS=EQ* VAR	14	.000069	.16	.9998	.030
Error	72	.002235			

R-squared = .913162.

Note. DS = Shape of conditional distributions; VAR = Conditional variance ratio; EQ = Sample size equality; Eta-squared = $SSH / (SSH + SSE)$. The sums of squares were obtained using the Type III sums of squares as defined in SAS PROC GLM procedure.

Table 13

Summary Statistics on the Type III Error Rates Yielded by the
Modified J-N Technique for Eight Shapes of the Conditional
Distribution

DS	n	Mean	LL	UL	Max	Q3	Med	Q1	Min
A	15	.0207	.0128	.0286	.0663	.0238	.0168	.0147	.0023
B	15	.0213	.0132	.0293	.0682	.0219	.0179	.0146	.0031
C	15	.0195	.0118	.0271	.0615	.0182	.0157	.0142	.0024
D	15	.0185	.0112	.0258	.0614	.0183	.0151	.0135	.0025
E	15	.0185	.0113	.0257	.0605	.0177	.0149	.0135	.0029
F	15	.0191	.0117	.0264	.0623	.0188	.0163	.0122	.0033
G	15	.0222	.0145	.0300	.0655	.0274	.0168	.0127	.0069
N	15	.0200	.0121	.0279	.0658	.0197	.0162	.0138	.0012

Note. DS = Conditional distribution shape; n = Number of observations; LL = Lower limit of the 95% confidence interval for the mean; UL = Upper limit of the 95% confidence interval for the mean; Max = Maximum value; Q3 = 75th percentile; Med = Median; Q1 = 25th percentile; Min = Minimum value.

Table 14

Summary Statistics on the Type III Error Rates Yielded by the
Modified J-N Technique for Three Conditional Variance Ratios
by Equal and Unequal Sample Sizes

EQ	VAR	n	Mean	LL	UL	Max	Q3	Med	Q1	Min
E	1	24	.0178	.0163	.0194	.0296	.0196	.0170	.0153	.0128
E	2	24	.0153	.0143	.0163	.0208	.0171	.0145	.0135	.0119
E	3	24	.0176	.0161	.0190	.0274	.0187	.0165	.0148	.0135
U	1	16	.0046	.0033	.0058	.0115	.0060	.0039	.0026	.0012
U	2	16	.0155	.0149	.0161	.0169	.0167	.0155	.0150	.0127
U	3	16	.0536	.0483	.0590	.0682	.0647	.0539	.0430	.0401

Note. EQ = Sample size equality (E = Equal, U = Unequal); VAR = Conditional variance ratio (1 = 18:54, 2 = 36:36, 3 = 54:18); n = Number of observations; LL = Lower limit of the 95% confidence interval for the mean; UL = Upper limit of the 95% confidence interval for the mean; Max = Maximum value; Q3 = 75th percentile; Med = Median; Q1 = 25th percentile; Min = Minimum value.